

RECEIVED
FEB 25 2002
TECH CENTER 1600/2900



1600

#12

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:02

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

P.S

ENTERED

```

3 <110> APPLICANT: MCKEON, FRANK
4     YANG, ANNIE
5     LODA, MASSIMO
6     SIGNORETTI, SABINA
7     CRUM, CHRISTOPHER
9 <120> TITLE OF INVENTION: CELL REGULATORY GENES, ENCODED PRODUCTS, AND USES
10    RELATED THERETO
12 <130> FILE REFERENCE: HNV-038.02
14 <140> CURRENT APPLICATION NUMBER: 09/538,106
15 <141> CURRENT FILING DATE: 2000-03-29
17 <150> PRIOR APPLICATION NUMBER: 09/174,493
18 <151> PRIOR FILING DATE: 1998-10-15
20 <150> PRIOR APPLICATION NUMBER: 60/087,216
21 <151> PRIOR FILING DATE: 1998-05-29
23 <150> PRIOR APPLICATION NUMBER: 60/062,076
24 <151> PRIOR FILING DATE: 1997-10-15
26 <160> NUMBER OF SEQ ID NOS: 53
28 <170> SOFTWARE: PatentIn Ver. 2.1
30 <210> SEQ ID NO: 1
31 <211> LENGTH: 1926
32 <212> TYPE: DNA
33 <213> ORGANISM: Homo sapiens
35 <220> FEATURE:
36 <221> NAME/KEY: CDS
37 <222> LOCATION: (1)..(1923)
39 <400> SEQUENCE: 1
40 atg tcc cag agc aca cag aca aat gaa ttc ctc agt cca gag gtt ttc      48
41 Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
42 1          5          10          15
44 cag cat atc tgg gat ttt ctg gaa cag cct ata tgt tca gtt cag ccc      96
45 Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
46          20          25          30
48 att gac ttg aac ttt gtg gat gaa cca tca gaa gat ggt gcg aca aac      144
49 Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
50          35          40          45
52 aag att gag att agc atg gac tgt atc cgc atg cag gac tcg gac ctg      192
53 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
54          50          55          60
56 agt gac ccc atg tgg cca cag tac acg aac ctg ggg ctc ctg aac agc      240
57 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
58 65          70          75          80
60 atg gac cag cag att cag aac ggc tcc tcg tcc acc agt ccc tat aac      288
61 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn

```

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:02

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

62	85	90	95	
64	aca gac cac gcg cag aac agc gtc acg gcg ccc tcg ccc tac gca cag	336		
65	Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln			
66	100	105	110	
68	ccc agc tcc acc ttc gat gct ctc tct cca tca ccc gcc atc ccc tcc	384		
69	Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser			
70	115	120	125	
72	aac acc gac tac cca ggc ccg cac agt ttc gac gtg tcc ttc cag cag	432		
73	Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln			
74	130	135	140	
76	tcg agc acc gcc aag tcg gcc acc tgg acg tat tcc act gaa ctg aag	480		
77	Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys			
78	145	150	155	160
80	aaa ctc tac tgc caa att gca aag aca tgc ccc atc cag atc aag gtg	528		
81	Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val			
82	165	170	175	
84	atg acc cca cct cct cag gga gct gtt atc cgc gcc atg cct gtc tac	576		
85	Met Thr Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr			
86	180	185	190	
88	aaa aaa gct gag cac gtc acg gag gtg gtg aag cgg tgc ccc aac cat	624		
89	Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His			
90	195	200	205	
92	gag ctg agc cgt gaa ttc aac gag gga cag att gcc cct cct agt cat	672		
93	Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His			
94	210	215	220	
96	ttg att cga gta gag ggg aac agc cat gcc cag tat gta gaa gat ccc	720		
97	Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro			
98	225	230	235	240
100	atc aca gga aga cag agt gtg ctg gta cct tat gag cca ccc cag gtt	768		
101	Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val			
102	245	250	255	
104	ggc act gaa ttc acg aca gtc ttg tac aat ttc atg tgt aac agc agt	816		
105	Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser			
106	260	265	270	
108	tgt gtt gga ggg atg aac cgc cgt cca att tta atc att gtt act ctg	864		
109	Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu			
110	275	280	285	
112	gaa acc aga gat ggg caa gtc ctg ggc cga cgc tgc ttt gag gcc cgg	912		
113	Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg			
114	290	295	300	
116	atc tgt gct tgc cca gga aga gac agg aag gcg gat gaa gat agc atc	960		
117	Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile			
118	305	310	315	320
120	aga aag cag caa gtt tcg gac agt aca aag aac ggt gat ggt acg aag	1008		
121	Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys			
122	325	330	335	
124	cgc ccg ttt cgt cag aac aca cat ggt atc cag atg aca tcc atc aag	1056		
125	Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys			
126	340	345	350	

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:02

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

```

128 aaa cga aga tcc cca gat gat gaa ctg tta tac tta cca gtg agg ggc 1104
129 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
130      355      360      365
132 cgt gag act tat gaa atg ctg ttg aag atc aaa gag tcc ctg gaa ctc 1152
133 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
134      370      375      380
136 atg cag tac ctt cct cag cac aca att gaa acg tac agg caa cag caa 1200
137 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
138 385      390      395      400
140 cag cag cag cac cag cac tta ctt cag aaa cag acc tca ata cag tct 1248
141 Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser
142      405      410      415
144 cca tct tca tat ggt aac agc tcc cca cct ctg aac aaa atg aac agc 1296
145 Pro Ser Ser Tyr Gly Asn Ser Ser Pro Pro Leu Asn Lys Met Asn Ser
146      420      425      430
148 atg aac aag ctg cct tct gtg agc cag ctt atc aac cct cag cag cgc 1344
149 Met Asn Lys Leu Pro Ser Val Ser Gln Leu Ile Asn Pro Gln Gln Arg
150      435      440      445
152 aac gcc ctc act cct aca acc att cct gat ggc atg gga gcc aac att 1392
153 Asn Ala Leu Thr Pro Thr Thr Ile Pro Asp Gly Met Gly Ala Asn Ile
154      450      455      460
156 ccc atg atg ggc acc cac atg cca atg gct gga gac atg aat gga ctc 1440
157 Pro Met Met Gly Thr His Met Pro Met Ala Gly Asp Met Asn Gly Leu
158 465      470      475      480
160 agc ccc acc cag gca ctc cct ccc cca ctc tcc atg cca tcc acc tcc 1488
161 Ser Pro Thr Gln Ala Leu Pro Pro Pro Leu Ser Met Pro Ser Thr Ser
162      485      490      495
164 cac tgc aca ccc cca cct ccg tat ccc aca gat tgc agc att gtc agt 1536
165 His Cys Thr Pro Pro Pro Pro Tyr Pro Thr Asp Cys Ser Ile Val Ser
166      500      505      510
168 ttc tta gcg agg ttg ggc tgt tca tca tgt ctg gac tat ttc acg acc 1584
169 Phe Leu Ala Arg Leu Gly Cys Ser Ser Cys Leu Asp Tyr Phe Thr Thr
170      515      520      525
172 cag ggg ctg acc acc atc tat cag att gag cat tac tcc atg gat gat 1632
173 Gln Gly Leu Thr Thr Ile Tyr Gln Ile Glu His Tyr Ser Met Asp Asp
174      530      535      540
176 ctg gca agt ctg aaa atc cct gag caa ttt cga cat gcg atc tgg aag 1680
177 Leu Ala Ser Leu Lys Ile Pro Glu Gln Phe Arg His Ala Ile Trp Lys
178 545      550      555      560
180 ggc atc ctg gac cac cgg cag ctc cac gaa ttc tcc tcc cct tct cat 1728
181 Gly Ile Leu Asp His Arg Gln Leu His Glu Phe Ser Ser Pro Ser His
182      565      570      575
184 ctc ctg cgg acc cca agc agt gcc tct aca gtc agt gtg ggc tcc agt 1776
185 Leu Leu Arg Thr Pro Ser Ser Ala Ser Thr Val Ser Val Gly Ser Ser
186      580      585      590
188 gag acc cgg ggt gag cgt gtt att gat gct gtg cga ttc acc ctc cgc 1824
189 Glu Thr Arg Gly Glu Arg Val Ile Asp Ala Val Arg Phe Thr Leu Arg
190      595      600      605
192 cag acc atc tct ttc cca ccc cga gat gag tgg aat gac ttc aac ttt 1872

```

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:02

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

```

193 Gln Thr Ile Ser Phe Pro Pro Arg Asp Glu Trp Asn Asp Phe Asn Phe
194      610                      615                      620
196 gac atg gat gct cgc cgc aat aag caa cag cgc atc aaa gag gag ggg      1920
197 Asp Met Asp Ala Arg Arg Asn Lys Gln Gln Arg Ile Lys Glu Glu Gly
198 625                      630                      635                      640
200 gag tga
201 Glu
204 <210> SEQ ID NO: 2
205 <211> LENGTH: 1551
206 <212> TYPE: DNA
207 <213> ORGANISM: Homo sapiens
209 <220> FEATURE:
210 <221> NAME/KEY: CDS
211 <222> LOCATION: (1)..(1548)
213 <400> SEQUENCE: 2
214 atg tcc cag agc aca cag aca aat gaa ttc ctc agt cca gag gtt ttc      48
215 Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
216 1      5      10      15
218 cag cat atc tgg gat ttt ctg gaa cag cct ata tgt tca gtt cag ccc      96
219 Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
220      20      25      30
222 att gac ttg aac ttt gtg gat gaa cca tca gaa gat ggt gcg aca aac      144
223 Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
224      35      40      45
226 aag att gag att agc atg gac tgt atc cgc atg cag gac tcg gac ctg      192
227 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
228      50      55      60
230 agt gac ccc atg tgg cca cag tac acg aac ctg ggg ctc ctg aac agc      240
231 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
232 65      70      75      80
234 atg gac cag cag att cag aac ggc tcc tcg tcc acc agt ccc tat aac      288
235 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn
236      85      90      95
238 aca gac cac gcg cag aac agc gtc acg gcg ccc tcg ccc tac gca cag      336
239 Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln
240      100      105      110
242 ccc agc tcc acc ttc gat gct ctc tct cca tca ccc gcc atc ccc tcc      384
243 Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
244      115      120      125
246 aac acc gac tac cca ggc ccg cac agt ttc gac gtg tcc ttc cag cag      432
247 Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
248      130      135      140
250 tcg agc acc gcc aag tcg gcc acc tgg acg tat tcc act gaa ctg aag      480
251 Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
252 145      150      155      160
254 aaa ctc tac tgc caa att gca aag aca tgc ccc atc cag atc aag gtg      528
255 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
256      165      170      175
258 atg acc cca cct cct cag gga gct gtt atc cgc gcc atg cct gtc tac      576

```

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:02

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

```

259 Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
260      180      185      190
262 aaa aaa gct gag cac gtc acg gag gtg gtg aag cgg tgc ccc aac cat 624
263 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
264      195      200      205
266 gag ctg agc cgt gaa ttc aac gag gga cag att gcc cct cct agt cat 672
267 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
268      210      215      220
270 ttg att cga gta gag ggg aac agc cat gcc cag tat gta gaa gat ccc 720
271 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
272 225      230      235      240
274 atc aca gga aga cag agt gtg ctg gta cct tat gag cca ccc cag gtt 768
275 Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
276      245      250      255
278 ggc act gaa ttc acg aca gtc ttg tac aat ttc atg tgt aac agc agt 816
279 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
280      260      265      270
282 tgt gtt gga ggg atg aac cgc cgt cca att tta atc att gtt act ctg 864
283 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
284      275      280      285
286 gaa acc aga gat ggg caa gtc ctg ggc cga cgc tgc ttt gag gcc cgg 912
287 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
288      290      295      300
290 atc tgt gct tgc cca gga aga gac agg aag gcg gat gaa gat agc atc 960
291 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
292 305      310      315      320
294 aga aag cag caa gtt tgc gac agt aca aag aac ggt gat ggt acg aag 1008
295 Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
296      325      330      335
298 cgc ccg ttt cgt cag aac aca cat ggt atc cag atg aca tcc atc aag 1056
299 Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
300      340      345      350
302 aaa cga aga tcc cca gat gat gaa ctg tta tac tta cca gtg agg ggc 1104
303 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
304      355      360      365
306 cgt gag act tat gaa atg ctg ttg aag atc aaa gag tcc ctg gaa ctc 1152
307 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
308      370      375      380
310 atg cag tac ctt cct cag cac aca att gaa acg tac agg caa cag caa 1200
311 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
312 385      390      395      400
314 cag cag cag cac cag cac tta ctt cag aaa cag acc tca ata cag tct 1248
315 Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser
316      405      410      415
318 cca tct tca tat ggt aac agc tcc cca cct ctg aac aaa atg aac agc 1296
319 Pro Ser Ser Tyr Gly Asn Ser Ser Pro Pro Leu Asn Lys Met Asn Ser
320      420      425      430
322 atg aac aag ctg cct tct gtg agc cag ctt atc aac cct cag cag cgc 1344
323 Met Asn Lys Leu Pro Ser Val Ser Gln Leu Ile Asn Pro Gln Gln Arg

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 02/11/2002

PATENT APPLICATION: US/09/538,106

TIME: 09:17:03

Input Set : A:\HNV-38.02 sequence listing.txt

Output Set: N:\CRF3\02112002\I538106.raw

L:3189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27